VALUE IN HEALTH 17 (2014) 406–415

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Modeling a Preference-Based Index for Two Condition-Specific Measures (Asthma and Overactive Bladder) Using a Nonparametric Bayesian Method

Samer A. Kharroubi, PhD[1,*], *John E. Brazier, PhD*[2], *Yaling Yang, PhD*[3]

[1]*Department of Mathematics, University of York, York, UK;* [2]*Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK;* [3]*Department of Primary Care Health Services, University of Oxford, Oxford, UK*

ABSTRACT

Background: Conventionally, parametric models were used for health state valuation data. Recently, researchers started to explore the use of nonparametric Bayesian methods in this area. **Objectives:** We present a nonparametric Bayesian model to estimate a preference-based index for two condition-specific five-dimensional health state classifications, one for asthma (five-dimensional Asthma Quality of Life Utility Index) and the other for overactive bladder (five-dimensional Overactive Bladder Quality of Life-Utility Index). **Methods:** Samples of 307 and 311 members of the UK general population valued 99 health states selected from a total of 3125 health states defined by each of the measures using the time trade-off technique. The article presents the results of the nonparametric model and compares it with the original model estimated using a conventional parametric random-effects model. The different methods are compared theoretically and in terms of empirical performance across the two data sets. It also reports the effect of respondent characteristics on health state valuations. **Results:** The nonparametric models were found to be better at predicting health state values within the estimation sample than without in terms of root mean square error and the patterns of standardized residuals. Some respondent characteristics were found to explain variation in health state values, but these did not have a significant effect on the health states values when estimates were adjusted for sample differences with the general population. **Conclusions:** The nonparametric Bayesian models are theoretically more appropriate than previously used parametric models and provide better utility estimates from the two condition-specific measures. Furthermore, the model is more flexible in estimating the effect of covariates.

Keywords: AQL-5D, asthma, covariates, nonparametric Bayesian methods, OAB-5D, overactive bladder, preference-based health measure, time trade-off.

Introduction

Preference-based measures of health-related quality of life are widely used to calculate quality-adjusted life-years (QALYs) for use in cost-effectiveness analyses. There are currently a number of generic preference-based measures, including the EuroQol five-dimensional (EQ-5D) questionnaire [1], health utilities index 2 (HUI2) and 3 [2,3], 15D [4,5], Assessment of Quality of Life [6], Quality of Well-Being [7], and the six-dimensional health state short form (derived from short-form 36 health survey) (SF-6D) [8]. Condition-specific preference-based measures have also been developed, such as for rhinitis [9], cancer [10], erectile dysfunction [11], overactive bladder [12], dementia [13], and urinary incontinence [14]. These measures have standardized multidimensional health state classifications with preference or utility weights elicited from a sample of the general population [15,16]. Their health state classifications, however, generate a large number of unique health states, so it is not feasible to obtain direct valuations for each health state. Models have to be estimated to predict the values for all states defined by the classification on the basis of direct valuations of a sample of states.

The distribution of health state values generated from time trade-off (TTO) and standard gamble (SG) tasks is typically skewed, truncated, noncontinuous, and hierarchical, and so present a major challenge for conventional statistical modeling [8]. Previous statistical models of these data have met with some success in the SF-6D [8], the EQ-5D questionnaire [17], and the HUI2 [18], but there are concerns with the size of the prediction errors and an apparent systematic pattern in the prediction errors (involving overprediction of the value of the poor health states and underprediction of the value of good health states). Furthermore, these methods are limited in the way they are able to model the effect of covariates on health state values. Covariates are modeled only in terms of their effect on the intercept, and these results in the intercept deviating from unity, which violates the requirement that full health equals one on the conventional full health-death scale used to estimate QALYs. It also makes the unrealistic assumption that

Note. Currently, the corresponding author is on a leave of absence from the University of York and is at Kuwait University, Department of Statistics & Operations Research, P.O.Box 5969, Safat 13060, Kuwait.

* *Address correspondence to*: Samer A. Kharroubi, Department of Mathematics, University of York, York YO24 3NB, UK.

E-mail: samer.kharroubi@york.ac.uk.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2014.05.002

covariates have the same effect regardless of the state (e.g., the effect is the same regardless of the severity of the state).

An alternative nonparametric Bayesian method for modeling health state preference data for handling these problems has been developed [19]. This method describes the intrinsic characteristics of individual health state valuation data in a way that is more theoretically appropriate than with previously used parametric models. It is also more flexible to take into account the effect of covariates. This method of modeling has been applied to SF-6D SG health state valuation data [20] and extended to address covariates [21]. It has also been applied to a generic measure for children, the HUI2 UK with SG valuation data [22]. In this article, we report on the application of this method to TTO data for two condition-specific preference-based measures: the five-dimensional Asthma Quality of Life Utility Index (AQL-5D) for asthma and the five-dimensional Overactive Bladder Quality of Life-Utility Index (OAB-5D) for overactive bladder. The results are compared with the conventional random-effect regression models for each data set. This article extends the method to two condition-specific measures and provides important evidence on the advantages of this approach to modeling TTO data.

The second section of this article briefly describes the AQL-5D and OAB-5D valuation studies and the data used in this article. The next section briefly sets out the parametric and nonparametric approaches for health state valuations. The fourth section presents the results from each approach and compares the models in terms of their ability to predict actual values and to estimate the effect of covariates. The last section concludes with a discussion of the results and their implication for future use of these measures and modeling work in this field.

The Valuation Survey and Data Set

The Health State Classifications

AQL-5D

The AQL-5D is a health state classification system [23] developed from the Asthma Quality of Life Questionnaire [24]. The dimensions of the AQL-5D are concern about asthma, shortness of breath, weather and pollution stimuli, sleep impact, and activity limitations (Table 1). The health state classification system has five dimensions, each with five levels of severity, with level 1 denoting no problems and level 5 indicating extreme problems. By selecting one level for each dimension, it is possible to define 3125 health states.

OAB-5D

The OAB-5D is a health state classification system [25] developed from the Overactive Bladder Questionnaire [12]. The dimensions of the OAB-5D are urge, urine loss, sleep, coping, and concern (Table 2). The health state classification system has the same structure as the AQL-5D, defining a total of 3125 health states.

The Surveys

The respondents were randomly sampled from the general population using the electoral register of names and address from within South Yorkshire, United Kingdom [11,12]. The TTO technique was chosen for eliciting preference values, which asks respondents to trade off between length of life and quality of life. The survey used the interviewer-administered TTO-prop method developed by the York Measurement and Valuation Health Group, which uses a "time board" as a visual aid [26]. This version of TTO was selected because it has been shown to be more reliable than a nonprops version [27]. Furthermore, it has been used to value the EQ-5D questionnaire.

Table 1 – Asthma quality-of-life classification (AQL-5D).

Concern

1. Feel concerned about having asthma none of the time
2. Feel concerned about having asthma a little or hardly any of the time
3. Feel concerned about having asthma some of the time
4. Feel concerned about having asthma most of the time
5. Feel concerned about having asthma all of the time

Short of breath

1. Feel short of breath as a result of asthma none of the time
2. Feel short of breath as a result of asthma a little or hardly any of the time
3. Feel short of breath as a result of asthma some of the time
4. Feel short of breath as a result of asthma most of the time
5. Feel short of breath as a result of asthma all of the time

Weather and pollution

1. Experience asthma symptoms as a result of air pollution none of the time
2. Experience asthma symptoms as a result of air pollution a little or hardly any of the time
3. Experience asthma symptoms as a result of air pollution some of the time
4. Experience asthma symptoms as a result of air pollution most of the time
5. Experience asthma symptoms as a result of air pollution all of the time

Sleep

1. Asthma interferes with getting a good night's sleep none of the time
2. Asthma interferes with getting a good night's sleep a little or hardly any of the time
3. Asthma interferes with getting a good night's sleep some of the time
4. Asthma interferes with getting a good night's sleep most of the time
5. Asthma interferes with getting a good night's sleep all of the time

Activities

1. Overall, not at all limited with all the activities done
2. Overall, a little limitation with all the activities done
3. Overall, moderate or some limitation with all the activities done
4. Overall, extremely or very limited with all the activities done
5. Overall, totally limited with all the activities done

AQL-5D, five-dimensional Asthma Quality of Life Utility Index.

The selection of health states was determined by the specification of the model to be estimated. In all 99 health states were selected out of the 3125 possible health states defined by the classification. The selection was on the basis of a balanced design, which ensured that any dimension-level (level λ of dimension δ) had an equal chance of being combined with all levels of the other dimensions. These 99 states were stratified into severity groups on the basis of their total level score across the dimensions (simply the sum of the levels), and then randomly allocated into 14 blocks, so that each block has 7 health

Table 2 – Overactive bladder quality-of-life classification system (OAB-5D).

Urge

1. Not at all bothered by an uncomfortable urge to urinate
2. Bothered by an uncomfortable urge to urinate a little bit or somewhat
3. Bothered by an uncomfortable urge to urinate quite a bit
4. Bothered by an uncomfortable urge to urinate a great deal
5. Bothered by an uncomfortable urge to urinate a very great deal

Urine loss

1. Not at all bothered by urine loss associated with a strong desire to urinate
2. Bothered by urine loss associated with a strong desire to urinate a little bit or somewhat
3. Bothered by urine loss associated with a strong desire to urinate quite a bit
4. Bothered by urine loss associated with a strong desire to urinate a great deal
5. Bothered by urine loss associated with a strong desire to urinate a very great deal

Sleep

1. Bladder symptoms interfered with your ability to get a good night's rest none of the time
2. Bladder symptoms interfered with your ability to get a good night's rest a little of the time
3. Bladder symptoms interfered with your ability to get a good night's rest some of the time
4. Bladder symptoms interfered with your ability to get a good night's rest a good bit or most of the time
5. Bladder symptoms interfered with your ability to get a good night's rest all of the time

Coping

1. Bladder symptoms caused you to plan "escape routes" to restrooms in public places none of the time
2. Bladder symptoms caused you to plan "escape routes" to restrooms in public places a little of the time
3. Bladder symptoms caused you to plan "escape routes" to restrooms in public places some of the time
4. Bladder symptoms caused you to plan "escape routes" to restrooms in public places a good bit or most of the time
5. Bladder symptoms interfered with your ability to get a good night's rest all of the time

Concern

1. Bladder symptoms caused you embarrassment none of the time
2. Bladder symptoms caused you embarrassment a little of the time
3. Bladder symptoms caused you embarrassment some of the time
4. Bladder symptoms caused you embarrassment a good bit or most of the time
5. Bladder symptoms caused you embarrassment all of the time

OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

states. This procedure ensured that all respondents, who were allocated 1 of the 14 blocks, received a set of states balanced in terms of severity and that each state was valued the same number of times apart from the pits state, which was valued by all respondents (see [11,12] for more details).

Table 3 – Characteristics of respondents in valuation surveys.

Characteristic	n (%)	
	AQL-5D	OAB-5D
Total	307	311
Age (y)		
18–25	34 (11.1)	37 (11.9)
26–35	57 (18.6)	57 (18.3)
36–45	61 (19.9)	61 (19.6)
46–55	50 (16.3)	51 (16.4)
56–65	45 (14.7)	45 (14.5)
>66	60 (19.5)	60 (19.3)
Female	168 (54.7)	160 (51.4)
Married or living with partner	214 (69.8)	217 (69.8)
Experienced serious illness		
In family	194 (63.4)	176 (56.6)
In themselves	94 (30.6)	94 (30.2)
Degree or equivalent	69 (22.5)	85 (27.3)
Education after 17 y	140 (45.6)	182 (58.5)
Renting property	64 (20.8)	63 (20.2)
Found valuation tasks in interview difficult		
Very difficult	24 (7.9)	13 (4.2)
Quite difficult	82 (26.7)	80 (25.9)
Neither difficult nor easy	52 (16.9)	70 (22.7)
Self-reported EQ-5D questionnaire scores		
Male, female	0.83, 0.84	0.88, 0.88

AQL-5D, five-dimensional Asthma Quality of Life Utility Index; EQ-5D, EuroQol five-dimensional; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

Respondents

Three hundred seven members of the public (response rate of 40%) in South Yorkshire (UK) were interviewed in the AQL-5D survey, and 311 people were interviewed in the OAB-5D survey (response rate of 26.7%). Table 3 shows that the two samples were very similar in terms of their sociodemographic composition. Among the respondents to the AQL-5D survey, 53 (17.3%) had asthma and in the OAB-5D survey, 27 (8.7%) reported experiencing symptoms of urge and 18 (5.8%) reported urine loss for at least some of the time. Overall self-reported health status using the EQ-5D questionnaire was very close to the UK EQ-5D questionnaire norms of 0.85 for females and 0.86 for males [28].

The Data Set

AQL-5D

There were 2455 TTO health state valuations generated by the 307 respondents from the interviews. The average number of TTO valuations per intermediate health state was 22 (range, 19–22) and the worst possible state (AQL-5D state 55555) was valued by every respondent (n = 307). Mean TTO health state values ranged from 0.39 to 0.94 and generally had fairly large SDs (around 0.2–0.4). The distribution of the values was negatively skewed.

OAB-5D

There were 2487 health state values generated by the 311 respondents. Each intermediate health state was valued 22

times using TTO (range, 17–29) and the worst possible state (OAB-5D 55555) was valued 310 times using TTO (one missing value). Mean TTO health state values ranged from 0.56 for the worst possible state, to 0.91 for state 13321, with an average SD of 0.28.

Modeling

The aim of modeling is to estimate health state utility values for all states defined by each classification. The utility associated with a health state is assumed to be a function of that state; hence, by estimating a relationship between the descriptive system and the observed values we can infer values for all states. In previous models of utility data, we have used parametric relationships with assumptions about functional form, but here we contrast this conventional parametric approach reported by Yang et al. [11,12] with a more flexible Bayesian nonparametric model.

A general model for health state valuations can be described by Kharroubi et al. [19]:

$$y_{ij} = f(\mathbf{X}_{ij}, \alpha_j) + \varepsilon_{ij}, \quad (1)$$

where for $i = 1, 2, \ldots, n_j$ and $j = 1, 2, \ldots, m$, $\mathbf{x}_{ij}$ is the ith health state valued by respondent j and the dependent variable y_{ij} is the TTO score given by respondent j for that health state. The general model has two sets of independent, zero-mean, random effect terms: ε_{ij} is a random error term associated with each observation and α_j is a term to allow for individual characteristics of respondent j.

The interpretation of $f(\mathbf{X}_{ij}, \alpha_j)$ is as the true indifference TTO value that respondent j has for health state $\mathbf{X}_{ij}$. The objective is to obtain a health state utility measure for the population as a whole, and this is generally taken to be the mean of the respondent-level health state utilities across the population. To account for different populations, it is possible to model α_j in terms of respondent-level covariates such as age, sex, or socioeconomic factors.

The Parametric Approach

Yang et al. [11,12] specify the following model for respondent j's health state utility:

$$f(\mathbf{X}_{ij}, \alpha_j) = \mu + \boldsymbol{\theta}' \mathbf{I}(\mathbf{X}_{ij}) + \alpha_j, \quad (2)$$

where μ and $\boldsymbol{\theta}$ denote unknown parameters and $\mathbf{I}(\mathbf{X}_{ij})$ is a vector of dummy explanatory variables. In the simplest, no-interactions, case of this model, $\mathbf{I}(\mathbf{X}_{ij})$ is a vector of terms $I_{\delta\lambda}(\mathbf{X}_{ij})$ for each level $\lambda > 1$ of dimension δ of the AQL-5D and/or OAB-5D. For example, $I_{32}(\mathbf{X}_{ij})$ denotes dimension $\delta = 3$ (weather and pollution in the AQL-5D; sleep in the OAB-5D), level $\lambda = 2$ (experience asthma symptoms as a result of air pollution a little or hardly any of the time in the AQL-5D; bladder symptoms interfered with your ability to get a good night's rest a little of the time in the OAB-5D). For any given health state $\mathbf{X}_{ij}$, $I_{\delta\lambda}(\mathbf{X}_{ij})$ is defined as follows:

$I_{\delta\lambda}(\mathbf{X}_{ij}) = 1$ if, for state $\mathbf{X}_{ij}$, dimension δ is at level λ.
$I_{\delta\lambda}(\mathbf{X}_{ij}) = 0$ if, for state $\mathbf{X}_{ij}$, dimension δ is not at level λ.

In all, there are 25 of these terms, with level $\lambda = 1$ acting as a baseline for each dimension. Hence, the intercept parameter μ represents the health state utility value for state (11111), and summing the coefficients $\theta_{\delta\lambda}$ of the "on" dummies and adding this to the intercept derives the value of any other state.

More generally, $\mathbf{I}(\mathbf{X}_{ij})$ can include additional dummy variables to account for interactions between the levels of different dimensions, and the model selected by Yang et al. [11,12] included one such term, Interaction dummy in the AQL-5D and N2_severe in the OAB-5D, each of which takes the value of 1 when two or more dimensions in a health state were greater than level 4 and 0 otherwise.

Estimation of this random-effects model is via generalized least squares or maximum likelihood. Because α_j has zero mean, the population health state utility for state $\mathbf{x}$ in this model is simply $\hat{\mu} + \hat{\boldsymbol{\theta}}' \mathbf{I}(\mathbf{X})$.

The Nonparametric Approach

The Bayesian statistical nonparametric model describes the characteristics of individual health state valuation data in such a way that for respondent j, the health state utility of state $\mathbf{X}_{ij}$ is

$$f(\mathbf{X}_{ij}, \alpha_j) = 1 - \exp(\alpha_j)\{1 - u(\mathbf{X}_{ij})\}. \quad (3)$$

Note that the individual respondent term α_j enters multiplicatively rather than additively as in Equation 2. The term $u(\mathbf{x})$ is the *median* health state utility of health state $\mathbf{x}$. (In the Kharroubi et al. [19] model, the distribution of α_j is normal, so it has zero median as well as zero mean, and the median of $\exp(\alpha_j)$ is therefore 1.) It is treated as an unknown function and in the Bayesian framework it therefore becomes a random variable. The prior distribution for $u(\mathbf{x})$ is

$$u(\mathbf{X}) \sim \mathrm{N}(\gamma + \boldsymbol{\beta}' X, \sigma^2), \quad (4)$$

Note that $\mathbf{x}$ is a vector comprising discrete levels on each of the five health dimensions. Note also that the mean function of Equation 4 represents a belief that the utility will be approximately linear and additive in different dimensions. This is not at all the same thing as model (2). Although model (2) imposes this linearity and additivity as a strict assumption about the utility function, model (4) simply expresses it as a prior expectation. (It is to be noted that linearity in each dimension is inappropriate because it is not reasonable to suppose that utility should drop by the same amount when we move from level 2 to 3 on a dimension as when we move from level 1 to 2. In the EQ-5D questionnaire, for instance, the levels in each dimension are of the form "no restriction" [in health or ability to function], "some restriction," "severe restriction," and there is no reason to suppose that the step from "none" to "some" has the same health impact as from "some" to "severe." Additivity may also be inappropriate, with some interaction to be expected. Linearity and additivity, however, can be overcome by models using a transformation function, a power or logarithmic function say. More on this is given in Kharroubi et al. [19].) The actual function is free to vary around this mean according to its multivariate normal distribution, and so it may take absolutely any form. It is in this sense that we describe our model as nonparametric, and we believe that this is another way in which our model is more realistic than that of Yang et al. [11,12]. See Kharroubi et al. [19] for more explanation of this part of the model.

Furthermore, the values of $u(\mathbf{X})$ and $u(\mathbf{X}')$ for two different states $\mathbf{x}$ and $\mathbf{X}'$ have a correlation $c(\mathbf{X}, \mathbf{X}')$ that decreases as the distance between $\mathbf{x}$ and $\mathbf{X}'$ increases. This is defined as

$$c(\mathbf{X}, \mathbf{X}') = \exp\left\{-\sum b_d (x_d - x'_d)^2\right\},$$

where for $d = 1, 2, \ldots, 5$, x_d and x'_d are the levels of dimension d in health states $\mathbf{x}$ and $\mathbf{X}'$, respectively, and b_d is a roughness parameter that controls how closely the true utility function is expected to adhere to a linear form in dimension d [19]. The effect of this function is to assert that if $\mathbf{x}$ and $\mathbf{X}'$ describe very similar health states (in the sense that their levels are the same or close in all dimensions) their utilities will be approximately the same, and so the preference function varies smoothly as the health state changes.

Note that the mean health state utility in Equation 3 is

$$\bar{u}(\mathbf{X}) = 1 - \bar{\alpha}\{1 - u(\mathbf{X})\},$$

where $\bar{\alpha}$ is the mean value of $\exp(\alpha)$ over the whole population. This will not in general be 1, and so the population (mean) health state utility is not the same as the median health state utility

$u(\mathbf{x})$. More details of the nonparametric modeling and evaluation of $\bar{\alpha}$ are given in Kharroubi et al. [19].

Note also that the variable α_j is allowed to vary because this captures the effect of respondent characteristics. Suppose that $\mathbf{t}_j$ is the vector of covariates for respondent j, we then propose

$$\alpha_j \sim N\left(\mathbf{t}_j'\theta, \tau^2\right)$$

where θ is the vector of coefficients for the covariates. Note here that $\mathbf{t}$'s are centered to ensure that they have zero means, and hence the value of $\exp(\alpha)$ for a typical person is 1.

The models and the programs to apply the Bayesian approach have been written in Matlab and are available on request.

Comparison of Models

The two models cannot be compared in terms of a simple table of coefficients as is the case in parametric models because the nonparametric model produces in effect a separate parameter for every one of the 3125 health states. Given the overall aim is to predict health state valuation, the best way to compare these models is via their predictive ability. This includes plots of predicted to actual values, calculations of the root mean squared error (RMSE), and plots of standardized residuals. These assessments are undertaken within the full estimation sample and in an out of sample random selection of 10 states by re-estimating the models using data sets excluding these 10 states.

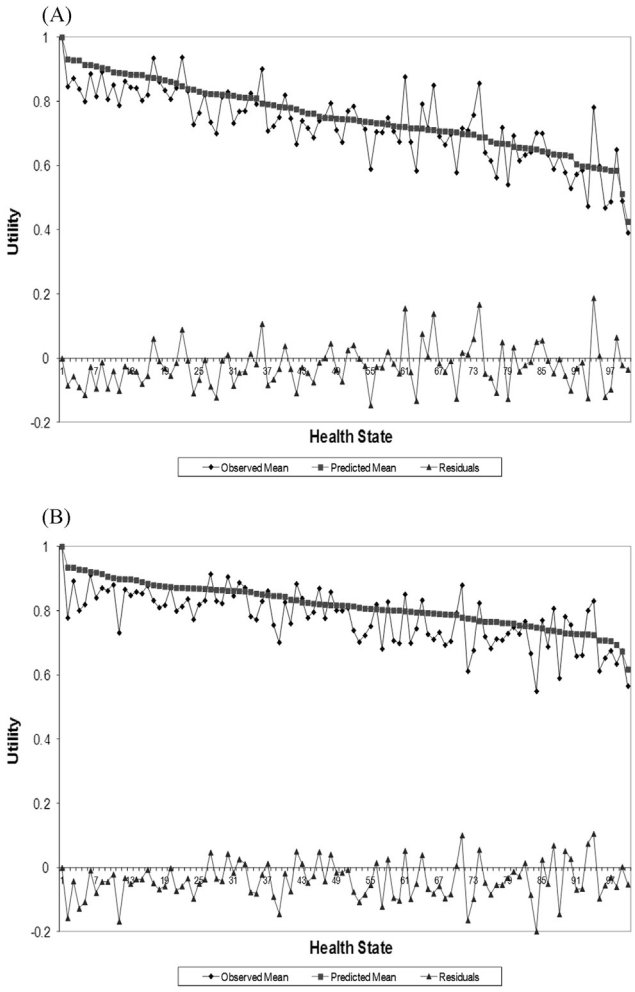

Fig. 1 – Sample mean and predicted health states valuations for the parametric model. (A) AQL-5D. (B) OAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

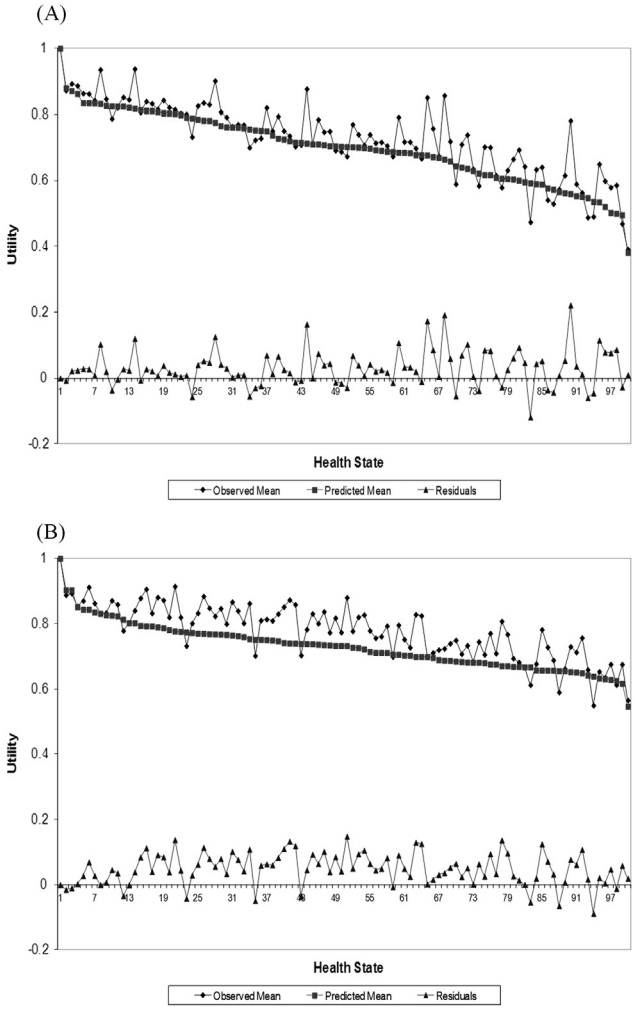

Fig. 2 – Sample mean and predicted health states valuations for the nonparametric model. (A) AQL-5D. (B) OAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

Results

The predictive ability of the two methods of modeling data is compared in Figures 1 and 2 for the two data sets. Figure 1 presents the predicted mean health state valuations (line marked with squares) for the parametric model (2), along with actual mean health state valuations (line marked with diamonds), with health states ordered by predicted health state values (Figs 1 and 2 have been plotted by ordering states in terms of their predicted values rather than observed values as presented in Yang et al. [11,12]. This is better on statistical grounds and easier to interpret.). The line marked with triangles represents the errors obtained by the difference between the two valuations. Figure 2 presents the same plots for the nonparametric model. These plots suggest that parametric models have a tendency to overpredict the value of the better states, whereas this does not seem to have been a problem for the nonparametric model. However, there is some suggestion that predictions are below those observed across the board in the OAB

Table 4 – Inference for a selection of health states.

A. AQL-5D

Health states	N	Observed mean	Mean ± SD		
			Nonparameteric posterior inference (no covariates)	Parameteric posterior inference	Nonparameteric posterior inference (covariates)
11233	21	0.8390	0.8115 ± 0.0283	0.9280 ± 0.0762	0.8125 ± 0.0283
12543	24	0.7475	0.7081 ± 0.0364	0.7801 ± 0.0724	0.7084 ± 0.0356
14225	25	0.7580	0.6720 ± 0.0366	0.6970 ± 0.0706	0.6719 ± 0.0358
15311	21	0.7643	0.7616 ± 0.0363	0.8306 ± 0.0744	0.7619 ± 0.0361
21223	25	0.8156	0.8082 ± 0.0270	0.9094 ± 0.0705	0.8088 ± 0.0265
23312	21	0.9376	0.8176 ± 0.0332	0.8476 ± 0.0765	0.8172 ± 0.0326
24422	19	0.7232	0.7527 ± 0.0272	0.7883 ± 0.0797	0.7532 ± 0.0263
25421	24	0.7921	0.7623 ± 0.0307	0.8100 ± 0.0719	0.7627 ± 0.0292
31215	22	0.7141	0.6933 ± 0.0351	0.7372 ± 0.0740	0.6959 ± 0.0339
32435	19	0.6916	0.7046 ± 0.0340	0.7075 ± 0.0799	0.7055 ± 0.0339
33511	21	0.9014	0.7761 ± 0.0400	0.7936 ± 0.0757	0.7760 ± 0.0394
34554	19	0.4905	0.5359 ± 0.0392	0.5116 ± 0.0803	0.5369 ± 0.0383
41125	25	0.6876	0.7033 ± 0.0325	0.7624 ± 0.0693	0.7033 ± 0.0324
42214	23	0.7400	0.6976 ± 0.0313	0.7530 ± 0.0730	0.6992 ± 0.0299
43234	23	0.7074	0.7000 ± 0.0309	0.7237 ± 0.0738	0.7010 ± 0.0292
45341	23	0.7704	0.7020 ± 0.0373	0.7444 ± 0.0742	0.7023 ± 0.0356
51454	23	0.5896	0.5533 ± 0.0415	0.6359 ± 0.0730	0.5550 ± 0.0395
52444	23	0.7009	0.6175 ± 0.0380	0.6449 ± 0.0730	0.6173 ± 0.0364
53532	25	0.7388	0.6363 ± 0.0367	0.7396 ± 0.0705	0.6353 ± 0.0359
55424	21	0.7819	0.5600 ± 0.0513	0.5940 ± 0.0767	0.5587 ± 0.0490
55555	307	0.3912	0.3811 ± 0.0377	0.4258 ± 0.0253	0.3828 ± 0.0345

B. OAB-5D

Health states	N	Observed mean	Mean ± SD		
			Nonparameteric posterior inference (no covariates)	Parameteric posterior inference	Nonparameteric posterior inference (covariates)
11233	21	0.8624	0.8350 ± 0.0243	0.9060 ± 0.0715	0.8356 ± 0.0247
12543	20	0.8305	0.7375 ± 0.0301	0.8645 ± 0.0736	0.7355 ± 0.0310
13514	20	0.8005	0.7365 ± 0.0311	0.8160 ± 0.0727	0.7355 ± 0.0325
15251	22	0.8368	0.7348 ± 0.0315	0.8703 ± 0.0699	0.7352 ± 0.0323
21113	25	0.8404	0.8018 ± 0.0262	0.9187 ± 0.0646	0.7999 ± 0.0274
23235	24	0.6617	0.6545 ± 0.0267	0.7269 ± 0.0676	0.6524 ± 0.0294
23312	19	0.9058	0.7930 ± 0.0268	0.8623 ± 0.0754	0.7922 ± 0.0277
24422	24	0.7017	0.7510 ± 0.0225	0.8460 ± 0.0674	0.7506 ± 0.0231
25421	20	0.8320	0.7925 ± 0.0245	0.8683 ± 0.0733	0.7917 ± 0.0254
31215	27	0.7519	0.7027 ± 0.0267	0.8056 ± 0.0632	0.7007 ± 0.0287
32435	24	0.5496	0.6381 ± 0.0296	0.7480 ± 0.0676	0.6375 ± 0.0318
33511	18	0.9144	0.7767 ± 0.0314	0.8667 ± 0.0766	0.7753 ± 0.0323
34554	24	0.5900	0.6548 ± 0.0293	0.7347 ± 0.0678	0.6553 ± 0.0308
35422	20	0.8270	0.7219 ± 0.0276	0.8442 ± 0.0739	0.7208 ± 0.0285
35453	23	0.7196	0.6889 ± 0.0283	0.7662 ± 0.0688	0.6883 ± 0.0305
41211	22	0.8809	0.7896 ± 0.0274	0.9017 ± 0.0695	0.7898 ± 0.0282
42325	24	0.7279	0.6565 ± 0.0284	0.7544 ± 0.0679	0.6544 ± 0.0301
44135	20	0.7705	0.6759 ± 0.0309	0.7453 ± 0.0732	0.6717 ± 0.0319
45143	21	0.7957	0.7054 ± 0.0312	0.8221 ± 0.0718	0.7043 ± 0.0331
45253	21	0.7671	0.6704 ± 0.0314	0.7525 ± 0.0721	0.6707 ± 0.0329
51451	22	0.8514	0.7413 ± 0.0334	0.7985 ± 0.0697	0.7428 ± 0.0339
52314	18	0.8333	0.8260 ± 0.0307	0.7938 ± 0.0774	0.8260 ± 0.0330
53525	25	0.6752	0.6167 ± 0.0355	0.6726 ± 0.0657	0.6124 ± 0.0392
54333	24	0.7071	0.6828 ± 0.0261	0.8011 ± 0.0686	0.6807 ± 0.0292
55555	310	0.5656	0.5464 ± 0.0300	0.6173 ± 0.0230	0.5448 ± 0.0331

AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

using the nonparametric model (Fig. 2B), though it is not systematic for the better or worst states, whereas the parametric model seems to predict higher utilities than those observed in the OAB (Fig. 1B).

Table 4 shows the inference for the mean health state utility values of a sample of 20 health states. These health states were selected at random from the 99 states valued in the sample to ensure a balance of mild, moderate, and severe states. For each state, Table 4 reports the observed sample mean health state utility and the predicted mean and SD for the population mean health state utility from both nonparametric and parametric models. It shows that across the two surveys the predictive performance of the nonparametric model is better than that of the parametric model overall, with an RMSE of 0.06 for the nonparametric model and 0.07 for the parametric model for both data sets.

No state is estimated as being worse than death in either data set. Important differences, however, can be seen between the models from Table 4. For AQL-5D, the parametric model estimates the health state utility for the pits state to be 0.4258, even though the observed average for this state is 0.3912, whereas the nonparametric model achieves a value of 0.3811. For OAB-5D, the parametric estimate was 0.6173 compared with an observed mean of 0.5656, but the nonparametric model gave 0.5464. The SDs of the predictions are larger for parametric models because they are based on the assumption that the preference function is additive in the various factors, apart from the interaction dummy term. The posterior SDs are smaller, and this is mainly because it is a model that allows uncertainty in the shape of the health state utility function.

The validity of the assumed models is examined in Figure 3, which shows a histogram of standardized residuals across all 2455 health state valuations for the parametric model, and Figure 4 shows the corresponding standardized residuals for the nonparametric model. According to these models, we would expect these to be approximately $N(0, 1)$. Figures 3 and 4 broadly support this, although there is some evidence of skewness, which is more obvious in Figure 3. This is not surprising, given the negative skewness in the original TTO data at the individual level. The degree of skewness, however, is probably not high enough to invalidate the analyses in both models, which assume normally distributed errors.

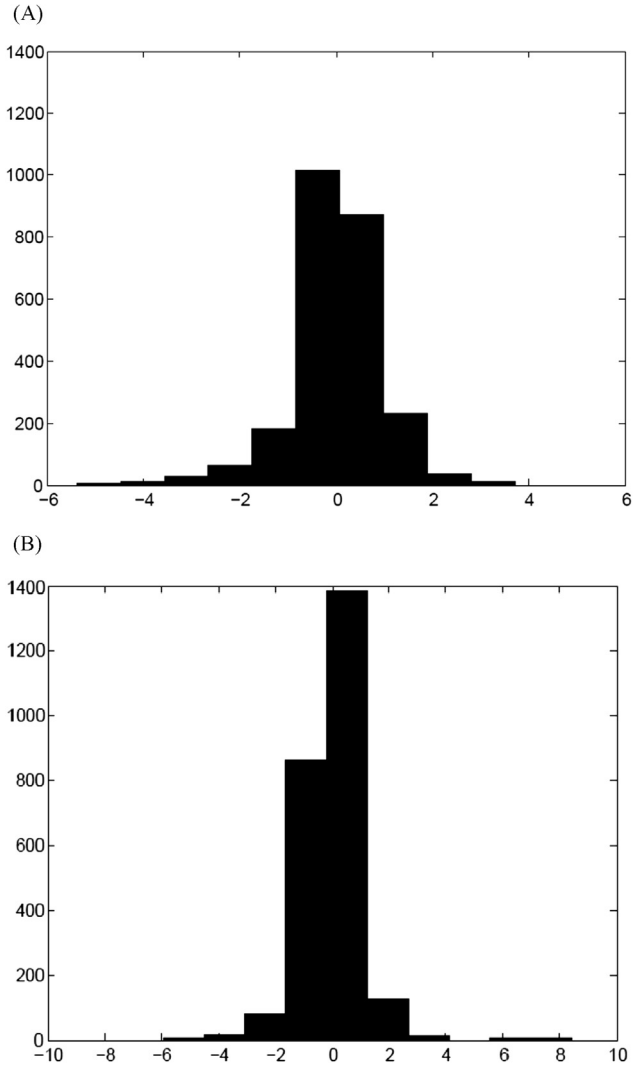

Fig. 3 – Standardized residuals for the parametric model for each of the 2455 individual health state valuations. (A) AQL-5D. (B) OAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

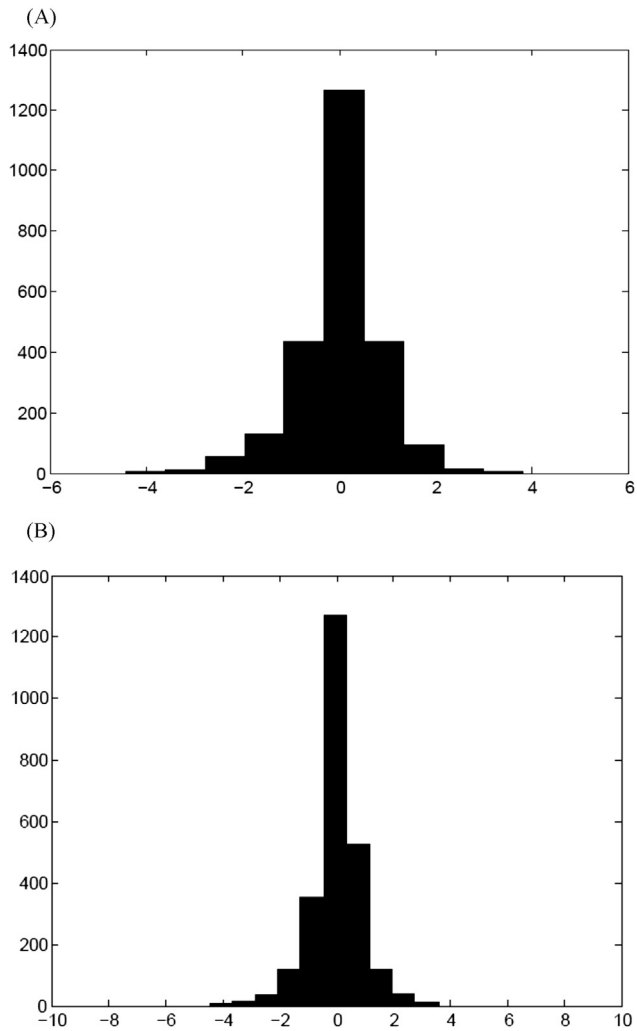

Fig. 4 – Standardized residuals for the nonparametric model for each of the 2455 individual health state valuation. (A) AQL-5D. (B) QAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

Table 5 – Out of sample predictions for 10 health states.

A. AQL-5D

Missing state	N	True sample mean	Mean ± SD	
			Nonparameteric posterior inference	Parameteric posterior inference
11445	22	0.6982	0.6665 ± 0.0788	0.6921 ± 0.0784
14225	25	0.7580	0.7789 ± 0.0628	0.6721 ± 0.0745
21113	25	0.8864	0.7941 ± 0.0693	0.9142 ± 0.0706
24422	19	0.7232	0.7494 ± 0.0582	0.7730 ± 0.0833
31143	23	0.8074	0.8177 ± 0.0543	0.8745 ± 0.0764
33245	22	0.7191	0.6850 ± 0.0564	0.6526 ± 0.0779
41112	23	0.8926	0.8784 ± 0.0531	0.9069 ± 0.0732
42542	19	0.7111	0.6756 ± 0.0628	0.7359 ± 0.0829
52141	20	0.8420	0.8149 ± 0.0815	0.8730 ± 0.0813
54333	23	0.7178	0.7120 ± 0.0589	0.7060 ± 0.0790

B. OAB-5D

Missing state	N	True sample mean	Mean ± SD	
			Nonparameteric posterior inference	Parameteric posterior inference
13321	20	0.9120	0.8388 ± 0.0441	0.9108 ± 0.0740
15331	23	0.8709	0.8141 ± 0.0423	0.9089 ± 0.0695
23534	25	0.7088	0.7679 ± 0.0388	0.7652 ± 0.0681
25425	22	0.7564	0.7560 ± 0.0435	0.7394 ± 0.0715
32441	21	0.8324	0.8301 ± 0.0411	0.8993 ± 0.0728
35422	20	0.8270	0.8365 ± 0.0402	0.8624 ± 0.0760
42245	24	0.6587	0.6717 ± 0.0402	0.7282 ± 0.0685
45532	20	0.7110	0.7304 ± 0.0508	0.8002 ± 0.0760
53242	20	0.7930	0.7510 ± 0.0447	0.7804 ± 0.0754
55521	23	0.7770	0.7530 ± 0.0528	0.8237 ± 0.0697

AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

A better test of the validity of the model is to investigate its ability to predict the values for states that have not been used in the estimation. To do this, 10 health states were removed randomly from the estimation data, and the models fitted on data for the remaining 89 states. The observed sample means for the 10 omitted states, together with their predicted mean and SD values from the parametric and nonparametric models estimated, are compared for the reduced data set in Table 5. It can be seen that the nonparametric model predicts the omitted data quite well and better than does the parametric model. Overall, the predictive performance of the nonparametric model is better than that of the parametric model, with RMSEs of 0.037 compared with 0.046, respectively, for the AQL-5D and 0.038 compared with 0.051 for the OAB-5D. The predicted SDs for both models are larger than those in Table 4, because the model in Table 4 is predicting the data on which it was estimated, whereas the model in Table 5 is predicting out of sample data. The parametric standard errors are larger than the nonparametric ones, primarily because the nonparametric analysis is able to make use of other evaluations by the same respondents to estimate their individual random effects, which the frequentist analysis cannot do.

The Q-Q plots of standardized predictive errors for the 10 health states sample means are presented in Figures 5 and 6. The dotted line on each is the fitted and corresponds to a reference line passing through first and third quartiles, which indicates whether the points are linear. The solid line corresponds to the theoretical $N(0,1)$ distribution. Figure 5 suggests that the parametric models are not well validated, particularly for the AOB-5D. In contrast, Figure 6 shows that nonparametric model predictions are well validated. Very similar results were obtained in five replications.

The analysis of covariates suggested that sex and age had some effect on AQL-5D health state values and experience of illness and age had some effect on OAB-5D values. Education did not have any discernible impact. As reported earlier, the age and sex distributions of respondents included in the two studies were different from those excluded and from the UK population as a whole. Therefore, we have examined the overall effect of adjusting for covariates on the values of the health states listed in Table 3. Actual UK age and sex distributions were taken from the UK census of 2011 [28]. The resultant posterior mean health state values were found to be almost identical to those that estimated the age and sex distributions from the valuation surveys, with differences of less than 0.005 in mean health state values.

Discussion

In this article, we have applied nonparametric methods to the existing AQL-5D and OAB-5D valuation data in an attempt to overcome some of the limitations observed when a parametric approach has been used. It extends methods used to model the SF-6D and HUI2 valuation data nonparametrically [19,22] to two condition-specific measures and applies them to TTO health state utility data. The Bayesian main-effects models estimated on two data sets have been compared with the conventional main effects model [11,12]. The flexibility of the nonparametric

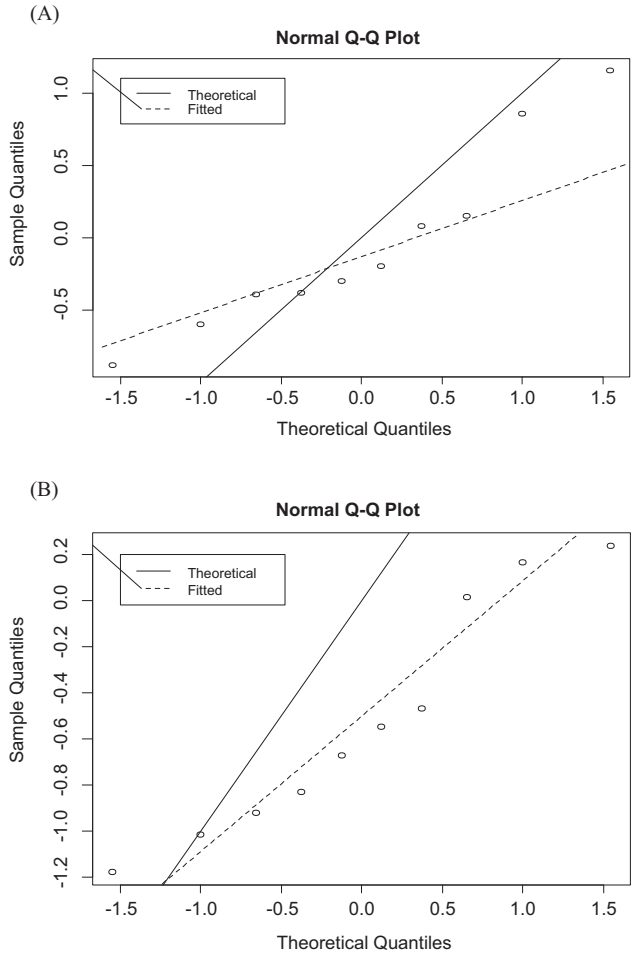

Fig. 5 – Q-Q plot of standardized predictive errors for the parametric model for the 10 out of sample health states. (A) AQL-5D. (B) OAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

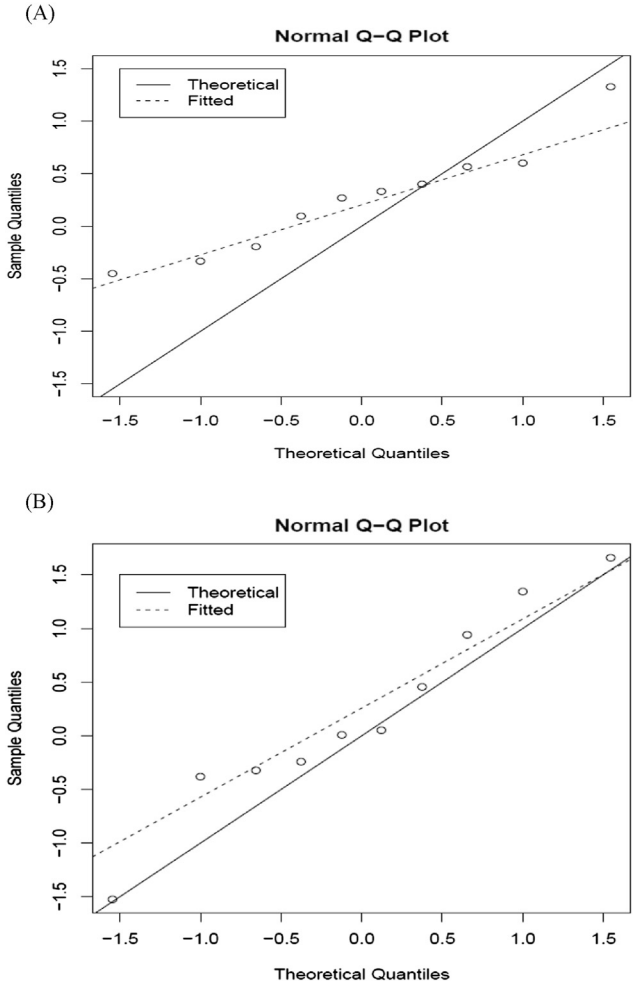

Fig. 6 – Q-Q plot of standardized predictive errors for the nonparametric model for the 10 out of sample health states. (A) AQL-5D. (B) OAB-5D. AQL-5D, five-dimensional Asthma Quality of Life Utility Index; OAB-5D, five-dimensional Overactive Bladder Quality of Life-Utility Index.

Bayesian method to examine the effect of covariates on health state values has also been used.

The nonparametric model is able to represent more accurately than the conventional parametric random-effects model the nature of individual respondent effects, the repeated measurements from each individual, and the skew distribution of individual valuations of a given state. The model also allows for respondent-level covariates to enter in a natural way as predictors of individual respondent effects.

As reported earlier, the covariates analysis showed that age and sex have strong effects on AQL-5D values and age and illness have a strong effect on OAQ-5D values. However, the overall effect of these background variables on the original posterior mean health state values after adjusting for age and sex were quite modest (typically 0.0045 or less). The descriptive content of the instruments seem to have been more important than these characteristics of the respondents examined in this article in explaining the variation in TTO values. Kharroubi at al. [21] drew a similar conclusion from their analysis of the SF-6D data.

The nonparametric model achieved improvements to the predictive ability in terms of a better RMSE and standardized predictive errors in the out of sample validation. Although the improvement in mean squared error is quite modest in this example, there are other important advantages over and above the mean difference. Our model produces

1. predictions that are much better behaved and do not violate logical consistency,
2. a utility function that better captures the lower utilities in the data for the poorest health states, and
3. posterior variances that correctly validate against the leftout data.

The extra flexibility of the nonparametric model and improved performance come at the expense of computational complexity and specialist software is needed to estimate the preference function (see [19] for details). The benefits are worth this extra given the expense of collecting good quality utility data from respondents. The programs developed for these applications can be readily adapted to other data sets with appropriate expertise.

The complexity of the modeling has little implication for those simply wishing to generate health state utility values from Asthma Quality of Life Questionnaire or Overactive Bladder Questionnaire data sets (e.g., from data collected in a clinical trial). The results can be readily applied using straightforward

Excel programs that have a procedure for undertaking a simple look-up procedure.

An alternative semi-parametric approach to health state valuation for handling these problems has been developed by Mendez et al. [29]. Their approach makes no assumption on the distribution of health state valuations, allows for an undetermined amount of heterogeneity in the estimates, and accommodates covariates in a flexible way. It is important to compare our approach to that applied by Mendez et al. in the future to make best use of health state valuation data.

This article presents nonparametric models to estimate a Bayesian preference-based index for two five-dimensional health state classifications, one for asthma (AQL-5D) and the other for overactive bladder (OAB-5D). It also reports on the effect of respondent characteristics on health state valuations. The nonparametric Bayesian models are theoretically more appropriate than previously used parametric models and provide better utility estimates from the two condition-specific measures. Furthermore, the model is more flexible in estimating the effect of covariates.

Acknowledgments

We thank Professor Tony O'Hagan for all his continual support, useful guidance, and invaluable insights during our time working on this article and Aki Tsuchiya and Tracey A. Young who were investigators in the original AQL-5D and OAB-5D valuation surveys.

Source of financial support: During the study, Yaling Yang was funded by a PhD studentship of the University of Sheffield. Professor John Brazier is funded by the Medical Research Council Health Service Research Collaboration.

REFERENCES

[1] Brooks R. EuroQol: the current state of play. Health Pol 1996;37:53–72.
[2] Torrance GW, Feeny DH, Furlong WJ, et al. Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. Med Care 1996;34:702–702.
[3] Feeny DH, Furlong WJ, Torrance GW, et al. Multi-attribute and single-attribute utility function for the Health Utility Index Mark 3 system. Med Care 2002;40:113–28.
[4] Sintonen H. The 15D-measure of health-related quality of life. I. Reliability, validity and sensitivity of its health state descriptive system (Working Paper 41). National Center for Health Program Evaluation, Melbourne, 1994.
[5] Sintonen H. The 15D-measure of health-related quality of life. II. Feasibility, reliability and validity of its valuation system (Working Paper 42). National Center for Health Program Evaluation, Melbourne, 1995.
[6] Hawthorne G, Richardson G, Atherton Day N. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Ann Med 2001;33:358–70.
[7] Kaplan RM, Anderson JP. A general health policy model: update and application. Health Serv Res 1988;23:203–35.
[8] Brazier JE, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002;21:271–92.
[9] Revicki DA, Leidy NK, Brennan-Diemer F, et al. Integrating patients' preferences into health outcomes assessment: the multiattribute asthma symptom utility index. Chest 1998;114:998–1007.
[10] Rowen D, Brazier J, Young T, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health 2011;14:721–31.
[11] Yang Y, Brazier JE, Tsuchiya A, Young T. Estimating a preference-based index for a 5-dimensional health state classification for asthma derived from the asthma quality of life questionnaire. Med Decis Making 2001;31:281–91.
[12] Yang Y, Brazier JE, Tsuchyia A, et al. Estimating a preference-based index from the Over Active Bladder questionnaire. Value Health 2009;12:159–66.
[13] Mulhern B, Rowen D, Brazier J, et al. Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. Health Technol Assess 2013;5:1–140:v–xv.
[14] Brazier JE, Czoski-Murray C, Roberts J, et al. Estimation of a preference-based index from a condition specific measure: the King's Health Questionnaire. Med Decis Making 2008;28:113–26.
[15] Brazier JE, Ratcliffe J, Tsuchiya A, Solomon J. Measuring and Valuing Health for Economic Evaluation. Oxford: Oxford University Press, 2007.
[16] Drummond MF, Sculpher M, O'Brien B, et al. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford Medical Publications, 2005.
[17] Dolan P. Modeling valuation for Euroqol health states. Med Care 1997;35:351–63.
[18] McCabe C, Stevens K, Roberts J, Brazier JE. Health state values for the HUI2 descriptive system: results from a UK Survey. Health Econ 2005;14:231–44.
[19] Kharroubi SA, O'Hagan A, Brazier JE. Estimating utilities from individual health state preference data: a nonparametric Bayesian approach. Appl Stat 2005;54:879–95.
[20] Kharroubi SA, Brazier J, O'Hagan A, Roberts J. Modelling SF-6D health state preference data using a nonparametric Bayesian method. J Health Econ 2007;26:597–612.
[21] Kharroubi SA, Brazier J, O'Hagan A, Roberts J. Modelling covariates for the SF-6D standard gamble health state preference data using a nonparametric Bayesian method. Soc Sci Med 2007;64:1242–52.
[22] Kharroubi SA, McCabe C. Modelling HUI 2 health state preference data using a nonparametric Bayesian method. Med Decis Making 2008;28:875–87.
[23] Young T, Yang Y, Brazier J, Tsuchyia A. Use of Rasch analysis in reducing a large condition specific instrument for preference valuation: the case of moving from AQLQ to AQL-5D. Med Decis Making 2011;31:195–210.
[24] Juniper EF, Guyatt GH, Ferrie PJ, et al. Measuring quality of life in asthma. Am Rev Resp Dis 1993;147:832–8.
[25] Young T, Yang Y, Brazier JE, et al. The first stage of developing preference-based measures: constructing a health state classification using Rasch analysis. Qual Life Res 2009;18:253–65.
[26] Gudex C. Time Trade-Off User Manual: Props and Self-Completion Methods. York, UK: Centre for Health Economics, University of York, 1994.
[27] Dolan P, Gudex C, Kind P. Valuing health states: a comparison of methods. J Health Econ 1996;2:209–32.
[28] The UK Office for National Statistics. Governments departments and devolved administrations. Available from: http://www.statistics.gov.uk/census. [Accessed January 29, 2013].
[29] Méndez I, Perpiñán JMA, Martínez FIS, Pérez JEM. Inverse probability weighted estimation of social tariffs: an illustration using the SF-6D value sets. J Health Econ 2011;30:1280–92.